# On environment-assisted capacities of quantum channels

Andreas Winter
*Department of Mathematics, University of Bristol, University Walk, Bristol BS8 1TW, U.K.*
*Email:* `a.j.winter@bris.ac.uk`
(Dated: 5th July 2005)

Following initial work by Gregoratti and Werner [J. Mod. Optics, vol. 50, no. 6&7, pp. 913-933, 2003 and `quant-ph/0403092`] and Hayden and King [`quant-ph/0409026`], we study the problem of the capacity of a quantum channel assisted by a "friendly (channel) environment" that can locally measure and communicate classical messages to the receiver.

Previous work [`quant-ph/0505038`] has yielded a capacity formula for the quantum capacity under this kind of help from the environment. Here we study the problem of the environment-assisted classical capacity, which exhibits a somewhat richer structure (at least, it seems to be the harder problem). There are several, presumably inequivalent, models of the permitted local operations and classical communications between receiver and environment: one-way, arbitrary, separable and PPT POVMs. In all these models, the task of decoding a message amounts to discriminating a set of possibly entangled states between the two receivers, by a class of operations under some sort of locality constraint.

After introducing the operational capacities outlined above, we show that a lower bound on the environment-assisted classical capacity is always half the logarithm of the input space dimension. Then we develop a few techniques to prove the existence of channels which meet this lower bound up to terms of much smaller order, even when PPT decoding measurements are allowed (assuming a certain superadditivity conjecture).

Keywords: entanglement of assistance, quantum error correction, feedback control, LOCC discrimination, PPT discrimination, additivity conjecture



## I. INTRODUCTION AND BACKGROUND

A noisy quantum channel is modelled universally as a completely positive and trace preserving (cptp) map

$$\mathcal{N} : A \longrightarrow B$$

between the algebras of observables $A = \mathcal{B}(\mathcal{H}_A)$ and $B = \mathcal{B}(\mathcal{H}_B)$, which we assume to be finite-dimensional throughout. It can always be presented as an isometry

$$U : \mathcal{H}_A \longrightarrow \mathcal{H}_B \otimes \mathcal{H}_C,$$

followed by the partial trace map $\mathrm{Tr}_C : B \otimes C \longrightarrow B$. This is the content of Stinespring's theorem [25], which also informs us that the isometry is unique up to unitaries on $\mathcal{H}_C$, which system is usually called the "channel environment". This means that associated with $\mathcal{N}$ there is a canonical "dual channel"

$$\overline{\mathcal{N}} : A \longrightarrow C,$$

defined as the above isometry $U$ followed by the other partial trace map $\mathrm{Tr}_B : B \otimes C \longrightarrow C$.

We shall here be interested in information transmission from A(lice) to B(ob) the channel, when assisted by the environment (Charlie), specifically in its information theoretic version of obtaining bounds on asymptotic rates.

An important case is where Bob and Charlie are allowed arbitrary local operations and classical communication (LOCC) to extract a classical encoded by Alice: the signals are then states $|\phi\rangle \in \mathcal{S} := U\mathcal{H}_A \subset \mathcal{H}_B \otimes \mathcal{H}_C$ in the image subspace $\mathcal{S}$ in the joint Bob-Charlie system, and the decoding task is to discriminate a set of these entangled states by measurements restricted by the LOCC constraint.

There are various important restrictions and relaxations of this model: we may insist on one-way LOCC from Charlie to Bob, or one-way LOCC from Bob to Charlie, or we may allow arbitrary LOCC. Since the class of all unrestricted LOCC operations, or even measurements, is notoriously hard to characterise, it is convenient for mathematical analysis to go to the larger class of separable POVMs, i.e. measurements whose POVM operators are sums of positive product operators, or in the even wider class of PPT (positive partial transpose) operators, as pioneered in Rains' work [20]: for $M = \sum_{ij,kl} M_{ij,kl} |ij\rangle\langle kl|$ (in an arbitrarily fixed basis), we demand $M^\Gamma := \sum_{ij,kl} M_{ij,kl} |il\rangle\langle kj| \geq 0$. It has been noticed before [4] that there are indeed separable POVMs which are not LOCC, and it is quite easy to see that there exist PPT POVMs which are not separable. Discriminating states via LOCC has become quite a large field, and here we can only collect a few pointers to the most significant papers (and references therein): Walgate et al. [27], Walgate and Hardy [28], Bennett et al. [4], Chefles [5], Ghosh et al. [9] and the more recent investigations by Badziąg et al. and Ghosh et al. [2], Nathanson [17] and Owari and Hayashi [19].

The structure of the paper is as follows: in the next section we will consider the problem of environment-assisted quantum capacity, and revisit the recently obtained capacity formula [24]. Then, in section III we introduce

the relevant notions of environment-assisted transmission codes and the corresponding capacities, and present various lower bounds. Section IV quotes a nontrivial upper bound on the LOCC-assisted classical capacity from [2], and presents an extension of it adapted to the more general class of PPT POVMs. Then, in section V we exhibit a class of examples for which the PPT-decoded classical capacity almost meets the general lower bound derived earlier, after which we conclude, highlighting a few open questions. An appendix quotes some technical results from the literature.

## II. QUANTUM CAPACITY WITH CLASSICAL HELPER IN THE ENVIRONMENT

Gregoratti and Werner [10] consider the channel model with helper in the environment, as outlined in the introductory section: an isometry $U$ from Alice's input system $A$ to the combination of Bob's output system $B$ and the environment $C$. Assume that the environment system may be measured and the classical results of the observation be forwarded to Bob — attempting to help him in error correcting quantum information sent from Alice.

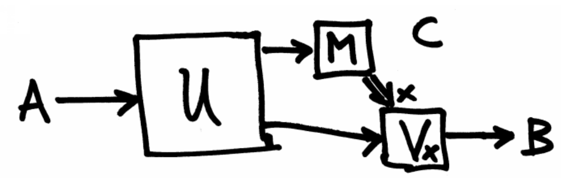


FIG. 1: Alice prepares an input to (many copies of) the isometry $U$, which gives part of the state to Bob and part to Charlie. The latter measures a POVM $M$ on his system and classically communicates his result $x$ to Bob, who executes a unitary $V_x$ depending on Charlie's message to recover Alice's input state.

We are interested, for this scenario, in the quantum transmission capacity from Alice to Bob, in the asymptotic limit of block coded information (and collectively measured environment). The setup is illustrated in figure 1. Formally, an environment-assisted quantum code (on block length $n$) is defined to consist of an encoder (cptp map) $\mathcal{E} : \mathcal{B}(\mathcal{H}) \longrightarrow \mathcal{B}(\mathcal{H}_A^{\otimes n})$, a POVM $(M_x)_x$ on $\mathcal{H}_C^{\otimes n}$ and cptp maps $\mathcal{R}_x : \mathcal{B}(\mathcal{H}_B^{\otimes n}) \longrightarrow \mathcal{B}(\mathcal{H})$; the idea is that Alice uses $\mathcal{E}$ to encode the quantum states she wants to send, Charlie performs the POVM $(M_x)_x$ and sends $x$ on to Bob who acts with the map $\mathcal{R}_x$ on the channel output. The overall dynamics $\mathcal{M} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ of this setup is

$$\mathcal{M}(\psi) = \sum_x \mathcal{R}_x\Big(\mathrm{Tr}_{C^n}\left[(U^{\otimes n}\mathcal{E}(\psi)U^{*\otimes n})(\mathbb{1}_B^{\otimes n} \otimes M_x)\right]\Big),$$

and we say that the code has error $\epsilon$ if for all $|\psi\rangle \in \mathcal{H}$, $\frac{1}{2}\|\psi - \mathcal{M}(\psi)\|_1 \leq \epsilon$. Incidentally, we will follow the convention of denoting a state vector always as a ket: e.g. $|\psi\rangle$, but its pure state density operator as $\psi = |\psi\rangle\langle\psi|$. Denoting by $M(n,\epsilon)$ the largest $\dim \mathcal{H}$ such that an environment-assisted quantum code on block length $n$ and with error $\epsilon$ exists, we can define the (optimistic/pessimistic) *environment-assisted quantum capacity* as

$$\inf_{\epsilon>0} \left( \limsup_{n\to\infty} / \liminf_{n\to\infty} \frac{1}{n} \log M(n,\epsilon) \right),$$

respectively.

In previous work by Smolin et al. [24], the following result was proved:

**Theorem 1 ([24], Thm. 8)** *The environment-assisted quantum capacity of the noisy channel $\mathcal{N} : A \longrightarrow B$ (both optimistically and pessimistically) is given by*

$$Q_A(\mathcal{N}) = \max_\rho \min\{S(\rho), S(\mathcal{N}(\rho))\}.$$

*The same capacity is obtained allowing unlimited LOCC between Alice, Bob and Charlie.* □

In particular, if the channel $\mathcal{N}$ is unital (i.e., preserving the identity) and $d_B \geq d_A$, then $Q_A(\mathcal{N}) = \log d_A$. In other words, all of the channel's input bandwidth can be corrected by looking at the environment. Since Gregoratti and Werner [10] have shown that perfect correction is possible if and only if the channel is a random mixture of isometries, this can be understood as saying that a unital channel $\mathcal{N}$ becomes, in the limit of many independent copies, almost a mixture of isometries. See [24] for a deeper discussion of this point.

## III. ENVIRONMENT- AND LOCC-ASSISTED CLASSICAL CAPACITIES

The isometry $U$ identifies $\mathcal{H}_A$ with the subspace $\mathcal{S} = U\mathcal{H}_A \subset \mathcal{H}_B \otimes \mathcal{H}_C$, so we can define a classical transmission code of blocklength $n$ as follows: it is a family $(\varphi_i, M_i)_{i=1}^N$ of pure states $|\varphi_i\rangle \in \mathcal{S}^{\otimes n}$ and a POVM $(M_i)_{i=1}^N$ on $\mathcal{H}_B^{\otimes n} \otimes \mathcal{H}_C^{\otimes n}$. We say that the code has error probability $\epsilon$ if for all $i$, $\mathrm{Tr}(\varphi_i M_i) \geq 1-\epsilon$. Some authors prefer the average error probability $\overline{\epsilon}$ as opposed to the maximal error we consider here: as in Shannon [23] it is easy to see that by expurgating the large-error signals, one can sacrifice a fraction $1/a$ of the messages and retain a set with maximal error $a\overline{\epsilon}$. Furthermore, with respect to the bipartition Bob-Charlie, we call the code

- *environment-assisted* if the POVM is implemented by one-way LOCC from Charlie to Bob;
- *environment-assisting* if the POVM is implemented by one-way LOCC from Bob to Charlie;
- *LOCC-assisted* if the POVM is implemented by some LOCC protocol;
- *separable-decoding* if the POVM is separable;

- *PPT-decoding* if the POVM consists of PPT operators.

The largest $N$ such that a code of blocklength $n$ and error probability $\epsilon$ exists under the above restrictions, are denoted $N_A^{\rightarrow}(n,\epsilon)$, $N_A^{\leftarrow}(n,\epsilon)$, $N_A^{\leftrightarrow}(n,\epsilon)$, $N_A^{\text{sep}}(n,\epsilon)$, $N_A^{\text{ppt}}(n,\epsilon)$, respectively.

Now we can define capacities in the usual way (cf. also the previous section): for example, the (one-way) environment-assisted classical capacity $C_A^{\rightarrow}(\mathcal{N})$ is given by

$$\inf_{\epsilon>0}\left(\limsup_{n\to\infty}\frac{1}{n}\log N_A^{\rightarrow}(n,\epsilon)\right)$$

in the optimistic version, and by

$$\inf_{\epsilon>0}\left(\liminf_{n\to\infty}\frac{1}{n}\log N_A^{\rightarrow}(n,\epsilon)\right)$$

in the pessimistic version, and likewise for $C_A^{\leftarrow}(\mathcal{N}) = C_A^{\rightarrow}(\overline{\mathcal{N}})$, $C_A^{\leftrightarrow}(\mathcal{N})$, $C_A^{\text{sep}}(\mathcal{N})$ and $C_A^{\text{ppt}}(\mathcal{N})$. We will not introduce special symbols for to distinguish optimistic and pessimistic capacities, but in this paper follow the convention that lower bounds on capacities are always proved pessimistically, and upper bounds optimistically.

Note that the models $\leftrightarrow$, sep and ppt are symmetric between Bob and Charlie; hence we denote, e.g. $C_A^{\text{ppt}}(\mathcal{N}) = C_A^{\text{ppt}}(U) = C_A^{\text{ppt}}(\mathcal{S})$, etc.

It should be obvious how to make the connection with the previously introduced capacity notions [10, 13]: for example, it is quite easy to see, using the well-known Fano inequality, that our capacity $C_A^{\rightarrow}(\mathcal{N})$ is the regularised "corrected Shannon capacity" $C_{\text{corr}}(\mathcal{N})$ of Hayden and King [13]:

$$C_A^{\rightarrow}(\mathcal{N}) = \lim_{n\to\infty}\frac{1}{n}C_{\text{corr}}\big(\mathcal{N}^{\otimes n}\big).$$

Clearly, we have the chain of inequalities

$$Q_A(\mathcal{N}) \leq C_A^{\rightarrow}(\mathcal{N}) \leq C_A^{\leftrightarrow}(\mathcal{N}) \leq C_A^{\text{ppt}}(\mathcal{N}) \leq \log d_A,$$

because every code to the left gives rise to or is itself immediately a code to the further right, and on the far right we have the input bandwidth, which is the capacity if B and C are permitted arbitrary joint operations.

For the formulation of the following general lower bound on $C_A^{\rightarrow}(\mathcal{N})$, let us introduce some notation: for a state $\rho$ on Alice's input system $A$, consider a generic purification $\phi$ on $A\otimes A$, and let $|\psi\rangle_{ABC} = (\mathbb{1}_A\otimes U)|\phi\rangle_{AA}$. Then denote the entropies of the reduced states of $\psi$ by referring to the subsystem(s) to which we restrict the state: e.g. $S(A) = S(\rho)$, $S(B) = S\big(\mathcal{N}(\rho)\big)$, $S(AB) = S\big(\operatorname{Tr}_C\psi\big) = S(C) = S\big(\overline{\mathcal{N}}(\rho)\big)$, etc. The quantum mutual information is formally defined as

$$I(A:B) = S(A)+S(B)-S(AB) = S(A)+S(B)-S(C).$$

For another state $\rho'$, we refer to the corresponding entropies by affixing primes: $S(A')$, $S(C')$, etc.

For example, theorem 1 implies that $C_A^{\rightarrow}(\mathcal{N}) \geq \min\{S(A),S(B)\}$ since one can always encode one bit in each qubit that is faithfully transmitted. Of course, we get by the same token $C_A^{\leftarrow}(\mathcal{N}) = C_A^{\rightarrow}(\overline{\mathcal{N}}) \geq \min\{S(A),S(C)\}$. By subadditivity of the entropy, $S(A) = S(BC) \leq S(B)+S(C)$, so the larger of $S(B)$ and $S(C)$ is at least $\frac{1}{2}S(A)$. Hence,

$$C_A^{\leftrightarrow}(U) \geq \max\big\{C_A^{\rightarrow}(\mathcal{N}), C_A^{\leftarrow}(\mathcal{N})\big\} \geq \frac{1}{2}\log d_A. \qquad (1)$$

Note that in general, by the above,

$$\begin{aligned} C_A^{\rightarrow}(\mathcal{N}) &\geq \min\big\{S(A),S(B)\big\} \\ &= \frac{1}{2}\big[S(A)+S(B)-|S(A)-S(B)|\big] \\ &\geq \frac{1}{2}\big[S(A)+S(B)-S(C)\big] = \frac{1}{2}I(A:B), \end{aligned}$$

the last line by the triangle inequality. Also, $\frac{1}{2}I(A:B) \geq \frac{1}{2}S(A)$ if $S(B) \geq S(C)$. And, if $S(A) \leq S(B)$, even $C_A^{\rightarrow}(\mathcal{N}) \geq S(A) \geq \frac{1}{2}S(A)$.

We shall now prove that also in the remaining case, $S(B) < S(C)$ and $S(B) < S(A)$, this lower bound holds, thus improving on eq. (1), in fact something a bit better. We will use the following recent result:

**Lemma 2 (State merging [16])** *Let $|\psi\rangle_{ABC}$ be a tripartite pure state with $S(A) = S(BC) < S(B)$. Then, for all $\epsilon > 0$ and all large enough $n$, there exists a measurement $(M_x)_x$ on $C^n$ and a family of isometries $V_x : \mathcal{H}_C^{\otimes n} \to \mathcal{H}_B^{\otimes n}\otimes\mathcal{H}_C^{\otimes n}$ such that*

$$\left\|\psi^{\otimes n} - \sum_x V_x \operatorname{Tr}_{C^n}\big[\psi^{\otimes n}(\mathbb{1}_{A^nB^n}\otimes M_x)\big]V_x^*\right\|_1 \leq \epsilon.$$

*If $S(A) \geq S(B)$, first sharing of $n\big(S(A)-S(B)\big)+o(n)$ ebits of entanglement creates a state which satisfies the above condition.* □

The protocol is based on time sharing between a block of length $k$ that is used to communicate $\sim k\big(S(C)-S(B)\big)$ bits from Alice to Charlie (who hands on the decoded message to Bob) and leaving $\sim kS(B)$ ebits between Bob and Charlie; and a block of length $\ell$ where Alice encodes $S(A')$ bits into a pure-state ensemble for $\rho'$ (and we assume $S(A') > S(B')$ here), Charlie merges his state with Bob's (lemma 2), using the previously extracted entanglement, so that Bob can read Alice's message, of $\sim \ell\, S(A')$ bits. On the first block we use random quantum coding for the channel $\overline{\mathcal{N}}$, see [6, 16] which justifies the transmission rate (of quantum information but we use an orthogonal basis in the code space to transmit classical information), and the remaining entanglement: see [6] for a description of the decoding via a unitary in Charlie's system, which separates the Alice's quantum message from the remaining entanglement. Per copy of the state, merging requires $S(A')-S(B')$ ebits and classical communication from Charlie to Bob [16], so we must

have $kS(B) \sim \ell\big(S(A') - S(B')\big)$. The rate is now the total information transmitted, $\sim k\big(S(C) - S(B)\big) + \ell\, S(A')$ bits, divided by the blocklength $k + \ell$. Thus we have proved:

**Theorem 3** *For the (one-way) environment-assisted classical capacity of the channel $\mathcal{N}$, and any input state $\rho$,*

$$C_A^{\rightarrow}(\mathcal{N}) \geq \begin{cases} S(A) & \text{if } S(A) \leq S(B), \\ \frac{1}{2} I(A:B) & \text{in general.} \end{cases} \tag{2}$$

*For input states $\rho$ such that $S(B) < S(C)$, and $\rho'$ such that $S(B') < S(A')$:*

$$C_A^{\rightarrow}(\mathcal{N}) \geq \frac{S(C) - S(B) + \frac{S(B)}{S(A') - S(B')} S(A')}{1 + \frac{S(B)}{S(A') - S(B')}}, \tag{3}$$

*so that for $\rho = \rho'$ with $S(B) < S(C)$ and $S(B) < S(A)$,*

$$C_A^{\rightarrow}(\mathcal{N}) \geq \left[1 - \frac{S(B)}{S(A)}\right] S(C) + \frac{S(B)}{S(A)} S(B). \tag{4}$$

□

**Corollary 4** *For every channel $\mathcal{N}$ and input state $\rho$,*

$$C_A^{\rightarrow}(\mathcal{N}) \geq \max\left\{\frac{1}{2} I(A:B), \frac{1}{2} S(A)\right\}, \tag{5}$$

*which for the maximally mixed input state $\rho = \frac{1}{d_A} \mathbb{1}$ gives that for all channels,*

$$C_A^{\rightarrow}(\mathcal{N}) \geq \frac{1}{2} \log d_A. \tag{6}$$

*Proof.* Note that $I(A:B) \geq S(A)$ if and only if $S(B) \geq S(C)$. Hence, we only have to show that for $\rho = \rho'$ with $S(B) < S(C)$ and $S(B) < S(A)$, the lower bound (4) in theorem 3 is at least as large as $\frac{1}{2} S(A)$:

$$\begin{aligned}
&\left[1 - \frac{S(B)}{S(A)}\right] S(C) + \frac{S(B)}{S(A)} S(B) \\
&\quad \geq \left[1 - \frac{S(B)}{S(B) + S(C)}\right] S(C) + \frac{S(B)}{S(B) + S(C)} S(B) \\
&\quad = \frac{S(B)^2 + S(C)^2}{S(B) + S(C)} \\
&\quad \geq \frac{1}{2}\big[S(B) + S(C)\big] \geq \frac{1}{2} S(A).
\end{aligned}$$

The first line comes from the subadditivity of entropy, $S(A) \leq S(B) + S(C)$, and using the assumption of $S(B) < S(C)$: substituting $S(B) + S(C)$ for $S(A)$ makes the weight of the smaller quantity smaller in the above convex combination. In the third line we use the arithmetic-geometric mean inequality, and subadditivity once more. □

The result that $C_A^{\rightarrow}(\mathcal{N}) \geq \frac{1}{2} \log d_A$ is somewhat reminiscent of an earlier observation by Fan [7]: that among the "standard" maximally entangled states in dimensions $d \times d$, any set of $\leq \sqrt{2d}$ is LOCC-distinguishable with certainty. Merging of quantum sources (lemma 2) gives is here an improvement in the asymptotic setting: consider any ensemble $\{p_i, \varphi_i\}$ of orthogonal entangled states on $BC$ such that $S(B) > S(BC)$ for the state $\rho = \sum_i p_i \varphi_i$. (For example, less than $d$ maximally entangled states in dimensions $d \times d$ with equal probabilities.) Then, for sufficiently many independent samples from the ensemble, Charlie can merge the unknown state from the ensemble with Bob's (at least with high fidelity and for a large-probability set of the ensemble), who then can distinguish them perfectly as they are orthogonal.

Another important lower bound, that is actually better than the above theorem and corollary for $d_A = 2, 3$, is proved in [13]: $C_A^{\rightarrow}(\mathcal{N}) \geq 1$ for every channel, which settles the capacity question for qubit input system.

## IV. AN UPPER BOUND ON THE PPT-DECODED CLASSICAL CAPACITY

In this section we will prove a general upper bound on the PPT-decoded classical capacity of a channel, and then demonstrate its usefulness by analysing a class of examples, in the following section.

Before we embark on this, we note that Badziąg et al. [2] have shown the following interesting bound:

**Proposition 5 ([2], Thm. 1)** *Consider an ensemble of pure states $|\varphi_i\rangle \in \mathcal{H}_B \otimes \mathcal{H}_C$, with probabilities $p_i$, and an LOCC-implemented POVM $(M_j)_j$. Then, with the joint distribution $\Pr\{X = i, Y = j\} = p_i \operatorname{Tr}(\varphi_i M_j)$ of random variables $X$ and $Y$, the Shannon mutual information is upper bounded as*

$$I(X:Y) \leq S(\rho_B) + S(\rho_C) - \overline{E}, \tag{7}$$

*where $\rho_{BC} = \sum_i p_i \varphi_i$ is the average state and $\rho_B$, $\rho_C$ are the reduced states, and $\overline{E} = \sum_i p_i E(\varphi_i)$ is the averaged pure state entanglement of the ensemble, and $E(\varphi) = S(\operatorname{Tr}_C \varphi)$.* □

This means that one obtains an upper bound on the "locally (rather: LOCC) accessible information". An interesting feature is that the term $\overline{E}$ vanishes if all states in the ensemble are products, but then in the example of [4] the above inequality is not tight. This motivates the conjecture that the above bound may be true for a much wider class of POVMs including all separable POVMs — indeed, as we will see at the end of this section, it holds true if the POVM is only PPT.

The following lemma is an adaptation of a result by Owari and Hayashi [19], whose is an elegant reformulation and proof of an insight by Nathanson [17], to the case of (small) error in the detection, not quite maximal entanglement, and PPT POVM elements:

**Lemma 6** *Consider Hilbert spaces $\mathcal{H}_B$ and $\mathcal{H}_C$ of dimensions $d_B \leq d_C$, respectively, and a pure state $|\varphi\rangle \in \mathcal{H}_B \otimes \mathcal{H}_C$ with $E(\varphi) \geq \log d_B - \Delta$. Then, for any PPT POVM element $M$ (i.e., $0 \leq M \leq \mathbb{1}$ and $M^\Gamma \geq 0$), such that $\operatorname{Tr}(\varphi M) \geq 1 - \epsilon$, and for every $K > 1$,*

$$\operatorname{Tr} M \geq \left(1 - \epsilon - \sqrt{2}\sqrt[4]{\Delta}\right) d_B, \tag{8}$$

$$\operatorname{Tr} M \geq \left(1 - \epsilon - \sqrt{\frac{\Delta + 1}{\log K}}\right) \frac{d_B}{K}. \tag{9}$$

*(The first bound is best for "small" $\Delta$, whereas the second will serve well in the regime of "large" $\Delta$.)*

*Proof.* For eq. (8) we observe that the condition $E(\varphi) = S(\operatorname{Tr}_C \varphi) \geq \log d_B - \Delta$ can be rewritten as

$$D\left(\operatorname{Tr}_C \varphi \middle\| \frac{1}{d_B}\mathbb{1}_B\right) \leq \Delta,$$

hence by Pinsker's inequality (lemma 16)

$$\frac{1}{2}\left\|\operatorname{Tr}_C \varphi - \frac{1}{d_B}\mathbb{1}_B\right\|_1 \leq \sqrt{\delta}.$$

Hence, using lemmas 14 and 15, there exists a maximally entangled state $\widehat{\varphi}$ (i.e. with $d_B$ Schmidt coefficients $1/d_B$) such that $F(\varphi, \widehat{\varphi}) \geq (1 - \sqrt{\delta})^2$, which implies (lemma 14 once more)

$$\frac{1}{2}\|\varphi - \widehat{\varphi}\|_1 \leq \sqrt{2}\sqrt[4]{\delta}.$$

From this get on the one hand

$$\operatorname{Tr}(\widehat{\varphi}M) \geq \operatorname{Tr}(\varphi M) - \sqrt{2}\sqrt[4]{\delta} \geq 1 - \epsilon - \sqrt{2}\sqrt[4]{\delta}.$$

On the other hand, using $M^\Gamma \geq 0$,

$$\begin{aligned}\operatorname{Tr}(\widehat{\varphi}M) &= \operatorname{Tr}\left(\widehat{\varphi}^\Gamma M^\Gamma\right) \leq \operatorname{Tr}\left(\left|\widehat{\varphi}^\Gamma\right| M^\Gamma\right) \\ &= \operatorname{Tr}\left(\frac{1}{d_B}\mathbb{1}\, M^\Gamma\right) = \frac{1}{d_B}\operatorname{Tr} M^\Gamma = \frac{1}{d_B}\operatorname{Tr} M.\end{aligned}$$

Here, we have used the modulus of an operator, $|A| = \sqrt{A^*A}$, and the fact that for a maximally entangled state, the partial transpose is the (unitary!) swap operator, divided by the Schmidt rank. This concludes the proof of eq. (8).

For eq. (9), let the Schmidt coefficients of $\varphi$ be denoted $\lambda_j$ ($j = 1, \ldots, d_B$), in decreasing order. We show first that

$$q := \sum\{\lambda_j : \lambda_j > K/d_B\} \leq \frac{\Delta + 1}{\log K}. \tag{10}$$

For this, assume that the first $L$ Schmidt coefficients $\lambda_j$ exceed $K/d_B$. From monotonicity of $H$ under majorisation (see Alberti and Uhlmann [1]) we see that the entropy of the distribution is maximised when $L = q\, d_B/K$ and the distribution has two flat sections: the first $L$ values are $q/L = K/d_B$, and the remaining $d_B - L$ values are $(1-q)/(d_B - L)$. (It is inessential for our argument that such $L$ may be non-integer: we only will overestimate the following entropy a little bit.) Now, this maximal entropy is

$$\begin{aligned}H &= H(q, 1-q) + q\log L + (1-q)\log(d_B - L) \\ &\geq E(\varphi) \geq \log d_B - \Delta.\end{aligned}$$

Rearranging this, using $H(q, 1-q) \leq 1$, and substituting $L/d_B = q/K$, this finally yields

$$\begin{aligned}\Delta + 1 &\geq -q\log\frac{q}{K} - (1-q)\log\left(1 - \frac{q}{K}\right) \\ &= q\log K - q\log q - (1-q)\log\left(1 - \frac{q}{K}\right) \\ &\geq q\log K,\end{aligned}$$

as claimed.

Now construct a pure state $\widetilde{\varphi}$ from $\varphi$ by removing all Schmidt coefficients exceeding $K/d_B$ (and normalising such as to obtain a unit vector): it is straightforward to check that $F(\varphi, \widetilde{\varphi}) = 1 - q$, with $q$ taken from eq. (10), hence (by lemma 14)

$$\frac{1}{2}\|\varphi - \widetilde{\varphi}\|_1 \leq \sqrt{q}.$$

From here we can proceed much as before: we first get

$$\operatorname{Tr}(\widetilde{\varphi}M) \geq \operatorname{Tr}(\varphi M) - \sqrt{q} \geq 1 - \epsilon - \sqrt{q}.$$

On the other hand, using $M^\Gamma \geq 0$ once more,

$$\begin{aligned}\operatorname{Tr}(\widetilde{\varphi}M) &= \operatorname{Tr}\left(\widetilde{\varphi}^\Gamma M^\Gamma\right) \leq \operatorname{Tr}\left(\left|\widetilde{\varphi}^\Gamma\right| M^\Gamma\right) \\ &\leq \operatorname{Tr}\left(\frac{K}{d_B}\mathbb{1}\, M^\Gamma\right) = \frac{K}{d_B}\operatorname{Tr} M^\Gamma = \frac{K}{d_B}\operatorname{Tr} M,\end{aligned}$$

which concludes the proof of eq. (9). □

**Remark 7** On completing the present manuscript its author has become aware of the recent paper [12]. It contains lower bounds similar to the above, for the error-free case, which are actually a bit better: for example, for a pure state $\varphi$ and separable/PPT POVM element $M$ with $\operatorname{Tr}(\varphi M) = 1$, it holds that $\log \operatorname{Tr} M \geq E(\varphi)$. It will be interesting to follow the development of the elegant techniques of [12] further, to deal with error probabilities.

**Theorem 8** *Let $(\varphi_i, M_i)_{i=1}^N$ be a code of pure states $|\varphi_i\rangle \in \mathcal{H}_B \otimes \mathcal{H}_C$, such that for all $i$, $E(\varphi_i) \geq \log d_B - \delta$, and PPT POVM $(M_i)_{i=1}^N$, with error probability $\leq \epsilon$. Then, if $\epsilon + \sqrt{2}\sqrt[4]{\delta} < 1$,*

$$N \leq \left(1 - \epsilon - \sqrt{2}\sqrt[4]{\delta}\right)^{-1} d_C.$$

*Proof.* Since by assumption all of the operators $M_i$ are PPT, we can use eq. (8) of lemma 6:

$$\text{for all } i,\ \operatorname{Tr} M_i \geq \left(1-\epsilon-\sqrt{2}\sqrt[4]{\delta}\right) d_B.$$

On the other hand, from the POVM condition that $\sum_{i=1}^N M_i \leq \mathbb{1}_{BC}$, we get that $\sum_{i=1}^N \operatorname{Tr} M_i \leq d_B d_C$, which yields the upper bound on $N$ as advertised. □

**Theorem 9** *Let $(\varphi_i, M_i)_{i=1}^N$ be a code of pure states $|\varphi_i\rangle \in \mathcal{H}_B \otimes \mathcal{H}_C$, such that for all $i$, $E(\varphi_i) \geq \log d_B - \Delta_i$, and PPT POVM $(M_i)_{i=1}^N$, with error probability $\leq \epsilon$. Then, for $\gamma > 1/(1-\epsilon)^2$,*

$$\begin{aligned} N &\leq \left(1-\epsilon-\sqrt{\frac{1}{\gamma}}\right)^{-1} \left(\frac{1}{N}\sum_{i=1}^N 2^{-\gamma(\Delta_i+1)}\right)^{-1} d_C \\ &\leq \left(1-\epsilon-\sqrt{\frac{1}{\gamma}}\right)^{-1} 2^{\sum_{i=1}^N \gamma(\Delta_i+1)/N} d_C. \end{aligned}$$

*Proof.* Since by assumption all of the operators $M_i$ are PPT, we can use eq. (9) of lemma 6: for the pair $(\varphi_i, M_i)$, we set $K_i = 2^{\gamma(\Delta_i+1)}$ and obtain

$$\text{for all } i,\ \operatorname{Tr} M_i \geq \left(1-\epsilon-\sqrt{\frac{1}{\gamma}}\right) 2^{-\gamma(\Delta_i+1)} d_B.$$

On the other hand, from the POVM condition that $\sum_{i=1}^N M_i \leq \mathbb{1}_{BC}$, we get that $\sum_{i=1}^N \operatorname{Tr} M_i \leq d_B d_C$, which yields the upper bound on $N$ as claimed; for the final upper bound we have to use the arithmetic-geometric mean inequality. □

**Corollary 10** *Let $U : \mathcal{H}_A \longrightarrow \mathcal{H}_B \otimes \mathcal{H}_C$ ($d_B \leq d_C$ without loss of generality), and assume for the subspace $\mathcal{S} = U\mathcal{H}_A \subset \mathcal{H}_B \otimes \mathcal{H}_C$, that for all $n$ and for all $|\varphi\rangle \in \mathcal{S}^{\otimes n}$, $E(\varphi) \geq n(\log d_B - \delta)$. Then,*

$$C_A^{\text{ppt}}(U) \leq \log d_C + \delta.$$

*Proof.* For a given blocklength $n$, consider a PPT-decoded code $(\varphi_i, M_i)_{i=1}^N$ of error $\leq \epsilon$ and rate $R = \frac{1}{n}\log N$.

We now use the previous theorem 9 with local dimensions $d_B^n$ and $d_C^n$, and $\Delta_i = n\delta$. This yields, for $\gamma > 1/(1-\epsilon)^2$,

$$N \leq \left(1-\epsilon-\sqrt{\frac{1}{\gamma}}\right)^{-1} 2^{\gamma(1+n\delta)} d_C^n.$$

For the rate this means

$$R \leq \log d_C + \gamma\delta + O\left(\frac{1}{n}\right),$$

and since in the limit $n \to \infty$, $\epsilon \to 0$ we can choose $\gamma$ arbitrarily close to 1, every asymptotically achievable rate is bounded above by $\log d_C + \delta$, as claimed. □

**Remark 11** The assumption of corollary 10 is widely believed to actually follow from the case $n = 1$. This is known as the superadditivity conjecture for the entanglement of formation [22]:

$$E_F(\rho_{B_1B_2C_1C_2}) \geq E_F(\rho_{B_1C_1}) + E_F(\rho_{B_2C_2}), \qquad (11)$$

where the entanglement of formation, $E_F$ [3], is the convex hull of the reduced state entropy function $E$.

Note that by assumption of $E(\varphi) \geq \log d_B - \delta$ for all $|\varphi\rangle \in \mathcal{S}$, every state $\rho$ supported on $\mathcal{S}$ has $E_F(\rho) \geq \log d_B - \delta$, hence by induction on $n$ we get $E(\varphi) \geq n(\log d_B - \delta)$ for all $|\varphi\rangle \in \mathcal{S}^{\otimes n}$.

**Remark 12** It should be obvious that the bound of Badziąg et al. [2], stated above as proposition 5, implies the bound of corollary 10 for the *LOCC-assisted capacity*:

$$C_A^{\leftrightarrow}(\mathcal{N}) \leq \log d_C + \delta.$$

Now we want to show that our theorems for PPT-decoders imply that the inequality (7) holds if $(M_j)_j$ is a PPT POVM.

*Proof (Sketch).* As before, the ensemble and the POVM give us random variables with joint distribution

$$\Pr\{X = i, Y = j\} = p_i \operatorname{Tr}(\varphi_i M_j).$$

Now, by Shannon's channel coding theorem [23], random coding on large block length $n$ gives, with high probability, a good code $\mathcal{C}$ of rate achieving $I(X:Y)$. In fact, since the codewords $I = i_1 \ldots i_n$ are chosen at random according to the distribution $p_I = p_{i_1} \cdots p_{i_n}$, most codewords will be typical, i.e., each letter $i$ occurs $\approx np_i$ times. Expurgating the untypical codewords we loose no rate asymptotically, but now all codewords can be assumed to be typical.

So, we have, for arbitrary $\eta > 0$ and for all sufficiently large $n$, a PPT-decoded code $(\Phi_I, D_I)_{I\in\mathcal{C}}$, with

$$|\Phi_I\rangle = |\varphi_{i_1}\rangle \otimes \cdots \otimes |\varphi_{i_n}\rangle$$

and PPT operators $D_I$, such that

$$\begin{aligned} \frac{1}{n}\log|\mathcal{C}| &\geq I(X:Y) - \eta, \\ \forall I \in \mathcal{C}\ \frac{1}{n}E(\Phi_I) &\geq \overline{E} - \eta, \end{aligned}$$

and error probability $\epsilon \to 0$ as $n \to \infty$.

Now we can further modify the code by using the typical subspace projectors $\Pi_B$ and $\Pi_C$ of $\rho_B^{\otimes n}$ and $\rho_C^{\otimes n}$, respectively [21]: create a new POVM (now on the tensor product of the two typical subspaces) with operators

$$D_I' := (\Pi_B \otimes \Pi_C) D_I (\Pi_B \otimes \Pi_C),$$

which is easily seen to inherit the PPT property from $(D_I)_I$. On the other hand (see [29]) this degrades the error probability only marginally, say increasing it to $2\epsilon$,

and the local dimensions of Bob and Charlie are now bounded by $2^{n(S(\rho_B)+\eta)}$ and $2^{n(S(\rho_C)+\eta)}$.

At this point we can finish, invoking theorem 9:

$$\begin{aligned}
I(X:Y) - \eta &\le \frac{1}{n}\log|\mathcal{C}| \\
&\le S(\rho_C) + \eta + \gamma\big(S(\rho_B) + \eta - \overline{E} + \eta\big) + O(1/n) \\
&\le S(\rho_B) + S(\rho_C) - \overline{E} \\
&\quad + (\gamma - 1)\log d_B + (2\gamma + 1)\eta + O(1/n).
\end{aligned}$$

Since $\eta > 0$ was arbitrary and also $\gamma > 1$ is arbitrary as the error probability $\epsilon \to 0$ and $n \to \infty$, we obtain the desired bound. ☐

## V. AN EXAMPLE ALMOST MEETING THE LOWER BOUND (6) ... MODULO ADDITIVITY CONJECTURE

It is clear that the upper bounds on $C_A^{\rm ppt}$ developed in the previous section are not very tight in general. In particular, for the bound of corollary 10 to be nontrivial, the dimension $d_A$ of the subspace $\mathcal{S}$ must be significantly larger than $d_C$ ($\ge d_B$).

Fortunately, we can use here the recently discovered existence of quite large subspaces in $d_B \times d_C$ which meet the requirements of theorem 8 and, assuming the universal validity of $E_F$-superadditivity (11), of corollary 10:

**Proposition 13 ([14], Thm. IV.1)** *Let $\mathcal{H}_B$ and $\mathcal{H}_C$ be quantum systems of dimension $d_B$ and $d_C$, respectively, for $d_C \ge d_B \ge 3$. Let $0 < \alpha < \log d_B$. Then there exists a subspace $\mathcal{S} \subset \mathcal{H}_B \otimes \mathcal{H}_C$ of dimension*

$$\left\lfloor d_B d_C \frac{\Gamma\alpha^{2.5}}{(\log d_B)^{2.5}} \right\rfloor$$

*such that all states $|\varphi\rangle \in \mathcal{S}$ have entanglement at least*

$$E(\varphi) = S(\varphi_A) \ge \log d_B - \frac{1}{\ln 2}\frac{d_B}{d_C} - \alpha,$$

*where $\Gamma$ is an absolute constant which may be chosen to be* $1/1753$. ☐

With $d_B = d_C = d$ and $\alpha = 20$ we are thus guaranteed a subspace $\mathcal{S} \subset \mathcal{H}_B \otimes \mathcal{H}_C$ of dimension $d_A = \left\lfloor d^2 \frac{1.0204}{(\log d)^{2.5}} \right\rfloor$, such that all states $|\varphi\rangle \in \mathcal{S}$ have entanglement $E(\varphi) \ge \log d - 21.5$. The channel $\mathcal{N}$ we now consider is simply the embedding $U$ of $\mathcal{H}_A = \mathcal{S}$ into the tensor product, followed by a partial trace over $C$. Of course this makes nontrivial sense only for rather large $d$ (namely $d \ge 128$, when $d_A$ starts becoming larger than $d$), which we silently assume from here on.

As mentioned a couple of times already, we will now assume the superadditivity conjecture (remark 11), which means that we will proceed under the assumption that

$$\text{for all } |\varphi\rangle \in \mathcal{S}^{\otimes n},\ E(\varphi) \ge n(\log d - 21.5).$$

Then corollary 10 gives us the bound

$$C_A^{\rm ppt}(U) \le \log d + 21.5 \le \frac{1}{2}\log d_A + 2.5\log\log d_A + 27.$$

The point here being that this comes close to the lower bound of theorem 3, up to a doubly logarithmic term and a (rather large) constant.

Finally, let us mention that using proposition 13 we can also produce an example catering to theorem 8: simply choose $d_B = d$, $d_C = \frac{2}{\delta \ln 2} d$ and $\alpha = \delta/2$.

## VI. DISCUSSION

We have shown some new lower and upper bounds on environment-assisted and PPT decoded capacities of quantum channels. In particular, we have shown that the environment-assisted classical capacity is always at least half the input bandwidth, and we have exhibited a class of examples which indicate that this factor of 1/2 is indeed attained in the worst case, even when the broader class of PPT decodings is permitted. This seems quite remarkable, as the lower bound is actually achieved some of the time by transmitting *quantum* information from Alice to Bob, and part of the time by transmitting quantum information partly to Charlie and partly to Bob (all of course with one-way LOCC help of Charlie to Bob).

In the process we have generalised a previously known bound on the locally accessible information to PPT POVMs; perhaps this will help clarifying the conceptual origin of such bounds (which in [2] is proved by going through a general LOCC protocol). It is however quite clear by simple examples that this upper bound cannot be optimal in general, even asymptotically and with coding; see also [15] which indicates that the upper bound cannot be in terms of local entropies and entanglement alone.

Our work still leaves wide open the problem of finding a formula for the assisted classical capacities $C_A^{\rightarrow}$ and $C_A^{\leftrightarrow}$. It seems that the main advance to be made is in trying to tighten the upper bounds on the locally accessible information. And of course we would like to narrow the gap between the lower bound and the worst-case upper bound for $C_A^{\rightarrow}$ and $C_A^{\rm ppt}$, and preferably so without resorting to unproven conjectures.

### Acknowledgments

Conversations with Chris King on the topics covered in this paper are gratefully acknowledged. The author is supported by the EU project RESQ (contract no. IST-2001-37559), by the U.K. Engineering and Physical Sciences Research Council's "IRC QIP", and a University of Bristol Research Fellowship.

## APPENDIX A: TECHNICAL RESULTS

**Lemma 14 (See [8])** *For two mixed states $\rho$, $\sigma$, the fidelity is $F(\rho,\sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1^2 = \left(\operatorname{Tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right)^2$, with the trace norm $\|A\|_1 = \operatorname{Tr}\sqrt{A^*A}$. Then,*

$$1 - \sqrt{F(\rho,\sigma)} \leq \frac{1}{2}\|\rho-\sigma\|_1 \leq \sqrt{1-F(\rho,\sigma)}.$$

**Lemma 15 (See [26])** *Let $\rho$, $\sigma$ be states on $\mathcal{H}$ and let $|\varphi\rangle, |\psi\rangle \in \mathcal{H}\otimes\mathcal{H}$ vary over purifications of $\rho, \sigma$, respectively. Then,*

$$F(\rho,\sigma) = \max_{\varphi,\psi} F(\varphi,\psi).$$

*Observe that for pure states, $F(\varphi,\psi) = \operatorname{Tr}\varphi\psi = |\langle\varphi|\psi\rangle|^2$.*

**Lemma 16 (Pinsker's inequality, see [18])** *For two arbitrary states $\rho$, $\sigma$, the relative entropy is defined as $D(\rho\|\sigma) = \operatorname{Tr}\big(\rho(\log\rho - \log\sigma)\big)$ [which may be $+\infty$ if the support of $\rho$ is not contained in that of $\sigma$]. Then,*

$$\left(\frac{1}{2}\|\rho-\sigma\|_1\right)^2 \leq D(\rho\|\sigma).$$

---


[1] P. M. Alberti, A. Uhlmann, *Stochasticity and partial order: doubly stochastic maps and unitary mixing*, Kluwer: Dordrecht-Boston, 1982.

[2] P. Badziąg, M. Horodecki, A. Sen(De), U. Sen, "Locally Accessible Information: How Much Can the Parties Gain by Cooperating?", Phys. Rev. Lett., vol. 91, no. 11, 117901, 2003. S. Ghosh, P. Joag, G. Kar, S. Kunkri, A. Roy, "Locally accessible information and distillation of entanglement", Phys. Rev. A, vol. 71, no. 2, 012321, 2005.

[3] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, W. K. Wootters, "Mixed-state entanglement and quantum error correction", Phys. Rev. A, vol. 54, no. 5, pp. 3824-3851, 1996.

[4] C. H. Bennett, D. P. DiVincenzo, C. A. Fuchs, T. Mor, E. Rains, P. W. Shor, J. A. Smolin, W. K. Wootters, "Quantum nonlocality without entanglement", Phys. Rev. A, vol. 59, no. 2, 1070-1091, 1999.

[5] A. Chefles, "Condition for unambiguous state discrimination using local operations and classical communication", Phys. Rev. A, vol. 69, no. 5, 050307(R), 2004.

[6] I. Devetak, "The Private Classical Capacity and Quantum Capacity of a Quantum Channel", IEEE Trans. Inf. Theory, vol. 51, no. 1, pp. 44-55, 2005.

[7] H. Fan, "Distinguishability and Indistinguishability by Local Operations and Classical Communication", Phys. Rev. Lett., vol. 92, no. 17, 177905, 2004.

[8] C. A. Fuchs, J. van de Graaf, "Cryptographic Distinguishability Measures for Quantum-Mechanical States", IEEE Trans. Inf. Theory, vol. 45, no. 4, pp. 1216-1227, 1999.

[9] S. Ghosh, G. Kar, A. Roy, A. Sen(De), U. Sen, "Distinguishability of Bell States", Phys. Rev. Lett., vol. 87, no. 27, 277902, 2001. S. Ghosh, G. Kar, A. Roy, D. Sarkar, "Distinguishability of maximally entangled states", Phys. Rev. A, vol. 70, no. 2, 022304, 2004.

[10] M. Gregoratti, R. F. Werner, "Quantum Lost and Found", J. Mod. Optics, vol. 50, no. 6&7, pp. 913-933, 2003.

[11] M. Gregoratti, R. F. Werner, "On quantum error correction by classical feedback in discrete time", e-print `quant-ph/0403092`, 2004.

[12] M. Hayashi, D. Markham, M. Murao, M. Owari, S. Virmani, "LOCC State Discrimination and Multipartite Entanglement Measures", e-print `quant-ph/0506170`, 2005.

[13] P. Hayden, C. King, "Correcting quantum channels by measuring the environment", e-print `quant-ph/0409026`, 2004.

[14] P. Hayden, D. W. Leung, A. Winter, "Aspects of generic entanglement", e-print `quant-ph/0407049`, 2004.

[15] M. Horodecki, A. Sen(De), U. Sen, K. Horodecki, "Local Indistinguishability: More Nonlocality with Less Entanglement" Phys. Rev. Lett., vol. 90, no. 4, 047902, 2003.

[16] M. Horodecki, J. Oppenheim, A. Winter, "Quantum information can be negative", e-print `quant-ph/0505062`, 2005. As "Partial quantum information" to appear in Nature.

[17] M. Nathanson, "Distinguishing Bipartitite Orthogonal States using LOCC: Best and Worst Cases", e-print `quant-ph/0411110`, 2004.

[18] M. Ohya, D. Petz, *Quantum Entropy and Its Use*, Springer Verlag, Berlin, 1993.

[19] M. Owari, M, Hayashi, "Local copying of orthogonal maximally entangled states and its relation to local discrimination", e-print `quant-ph/0411143`, 2004.

[20] E. M. Rains, "A Semidefinite Program for Distillable Entanglement", IEEE Trans. Inf. Theory, vol. 47, no. 7, pp. 2921-2933, 2001.

[21] B. Schumacher, "Quantum Coding", Phys. Rev. A, vol. 51, no. 4, pp. 2738-2747, 1995. R. Jozsa, B. Schumacher, "A new proof of the quantum noiseless coding theorem", J. Mod. Optics, vol. 41, no. 12, pp. 2343-2349, 1994.

[22] K. G. H. Vollbrecht, R. F. Werner, "Entanglement measures under symmetry", Phys. Rev. A, vol. 64, no. 6, 062307, 2001. K. Matsumoto, T. Shimono, A. Winter, "Remarks on Additivity of the Holevo Channel Capacity and of the Entanglement of Formation", Comm. Math. Phys., vol. 246, no. 3, pp. 427-442, 2004. P. W. Shor, "Equivalence of Additivity Questions in Quantum Information Theory", Comm. Math. Phys., vol. 246, no. 3, pp. 453-472, 2004. A. A. Pomeransky, "Strong superadditivity of the entanglement of formation follows from its additivity", Phys. Rev. A, vol. 68, no. 3, 032317, 2003.

[23] C. E. Shannon, "A mathematical theory of communication", Bell. Syst. Tech. J., vol. 27, pp. 379-423 and 623-656, 1948.

[24] J. A. Smolin, F. Verstraete, A. Winter, "Entanglement of assistance and multipartite state distillation", e-print `quant-ph/0505038`, 2005.

[25] W. F. Stinespring, "Positive functions on C$^*$-algebras",

Proc. Amer. Math. Soc., vol. 6, pp. 211-216, 1955.

[26] A. Uhlmann, "The 'transition probability' in the state space of a $*$-algebra", Rep. Math. Phys., vol. 9, pp. 273-279, 1976. R. Jozsa, "Fidelity for mixed quantum states", J. Mod. Optics, vol. 41, no. 12, pp. 2315-2323, 1994.

[27] J. Walgate, A. Short, L. Hardy, V. Vedral, "Local Distinguishability of Multipartite Orthogonal Quantum States, Phys. Rev. Lett., vol. 85, no. 23, pp. 4972-4975, 2000.

[28] J. Walgate, L. Hardy, "Nonlocality, Asymmetry, and Distinguishing Bipartite States", Phys. Rev. Lett., vol. 89, no. 14, 147901, 2002.

[29] A. Winter, "Coding theorem and strong converse for quantum channels", IEEE Trans. Inf. Theory, vol. 45, no. 7, pp. 2481-2485, 1999.